Bayesian Learning for Individualized Treatment Decisions in Personalized Medicine

Pamela Massiel Chiroque Solano, University of Regensburg

Personalized medicine seeks to identify the most appropriate treatment for an individual patient rather than relying solely on average treatment effects observed in clinical trials. However, translating heterogeneous patient information into reliable treatment recommendations remains a major statistical challenge.

This work presents a Bayesian statistical framework for estimating Predicted Individual Treatment Effects (PITE) by integrating evidence from clinical data. The approach captures complex treatment-response patterns while maintaining interpretability, allowing clinicians to understand how patient characteristics contribute to treatment recommendations.

Unlike conventional predictive models that focus primarily on prediction accuracy, the proposed framework explicitly supports treatment decision-making by combining individualized treatment effect estimation with decision-theoretic principles. It integrates high-dimensional and heterogeneous data while quantifying uncertainty through Bayesian Additive Regression Trees (BART). By producing transparent, patient-specific evidence, it supports clinicians in selecting treatments better aligned with individual patient profiles.

Applied to Alcohol Use Disorder (AUD), this work demonstrates how Bayesian statistical learning can inform individualized treatment decisions by integrating predictive modelling with principled decision support.

metaextractoR: An R Package for Data Extraction using LLMs for meta-analysis

Danyang Dai, University of Queensland

Systematic reviews and meta-analyses are central to evidence-based medicine, but data extraction is time-consuming and error-prone. Although Large Language Models (LLMs) show promise for automating this process, their adoption remains limited. We introduce metaextractoR, an open-source R package that streamlines data extraction by integrating open-source LLMs via the Ollama framework. The package supports automated, structured, and reproducible workflows through interactive Shiny applications, aligned with Cochrane guidelines and veridical data science principles.

metaextractoR includes three modular apps: (1) a manual extraction interface, (2) tools for prompt engineering and model selection, and (3) validation of LLM outputs. Together, these enable a human-in-the-loop approach, where each abstract is double-extracted by both a researcher and an LLM, mirroring best-practice recommendations. The package runs entirely on local machines, eliminating the need for APIs and ensuring data privacy. Comprehensive logging of prompts and outputs enhances transparency and reproducibility. By lowering technical barriers, metaextractoR enables more efficient, reliable, and transparent evidence synthesis.

Innovative trial designs: When not to use it

Aritra Mukherjee, Newcastle University

Over the last few years, the use of innovative trial designs has gained much momentum in practice. These designs encompass adaptive designs, platform trials, master protocols, among others. They offer greater flexibility without undermining trial integrity, often improving efficiency, patient benefit, and reducing the time to identify effective treatments. However, these advantages may be adversely affected by delays in observing the primary outcome variable. In the presence of such delays, the choice is between (a) pausing recruitment until sufficient data are accrued for the interim analysis, leading to longer times to identify effective treatment arms; or (b) continuing to recruit patients, which may result in many participants who do not benefit from the interim analysis. Little work has investigated the extent of outcome delay that results in the realised efficiency gains of these trial designs meeting the planned gains.

We assess the impact of outcome delay on the expected efficiency gains of different adaptive designs and platform trials by estimating the number of pipeline patients recruited under the assumption that recruitment is not paused while awaiting treatment outcomes.

The results indicate that if outcome delay is not considered at the planning stage, much of the expected efficiency gains may be lost. In light of these findings, we propose guidelines for clinicians when planning innovative trial designs in the presence of delayed outcomes.

Menopause-related heterogeneity in lipoprotein(a) effect on atherosclerotic cardiovascular disease using Mendelian Randomization

Daianna Gonzalez-Padilla, MRC Biostatistics Unit, University of Cambridge

Atherosclerotic cardiovascular disease (ASCVD) risk rises in women after menopause, making postmenopausal women a key prevention group. Lipoprotein(a) [Lp(a)] is a promising therapeutic target as it contributes to ASCVD risk and increases after menopause. However, whether its therapeutic potential could vary across menopausal stages remains unclear.
Mendelian randomization (MR) uses genetic variants associated with an exposure to estimate its causal effect on an outcome. While MR is well suited to assess the role of Lp(a) in ASCVD, stratifying analyses by menopausal status may induce collider bias since menopause can be influenced by Lp(a)-related variants and ASCVD risk factors. In contrast, conditioning on sex and age avoids this bias, as neither is affected by Lp(a) genetics or cardiovascular risk factors.
Using the UK Biobank, we evaluated the effect of Lp(a) on coronary and peripheral artery disease across sex and age strata. Lp(a) was associated with increased risk in both women and men, with no evidence of female-specific divergence across age strata, supporting a consistent causal contribution of Lp(a) to ASCVD risk before and after menopause. More broadly, we propose sex- and age-stratified MR frameworks for studying menopause–related heterogeneity in cardiovascular aetiology.